PROBABILIDAD Y ESTADÍSTICA
EXAMEN, 16 de diciembre de 2010

DATOS DEL ESTUDIANTE

No. de Examen	Nombre y Apellido	Cédula

- **La duración del examen es de 4 horas.**
- **Publicación de resultados: miércoles 22 de diciembre a las 20:00 horas.**
- **Muestra de exámenes: jueves 23 de diciembre a las 14:00 horas.**

Problema 1 (30 puntos)

Se tiene una cajonera con 3 cajones, en el primer cajón hay 3 medias, en el segundo 4 medias y en el tercero 5 medias. En cada cajón exactamente 2 de las medias son negras y el resto blancas. El experimento consiste en elegir un cajón al azar y luego comenzar a sacar medias (del cajón seleccionado y también al azar), sin devolverlas, hasta obtener 2 medias negras.

1. Hallar la probabilidad de obtener la segunda media negra a la quinta vez de haberlo intentado.
2. Sea X la variable aleatoria que indica cuántas medias se sacaron hasta obtener la segunda media negra. Hallar el recorrido, la función de probabilidad y la función de distribución de X.
3. Repetir la parte 1., pero ahora cada vez que se saca una media se la devuelve al cajón.

Justificar en todos los casos, e indicar si se encuentran distribuciones conocidas.

Problema 2 (30 puntos)

Sean X con distribución binomial de parámetros n y $\frac{1}{n}$ ($X \sim \text{Bin}(n, 1/n)$) e Y con distribución de Poisson de parámetro λ ($Y \sim \mathcal{P}(\lambda)$) variables aleatorias independientes. Se considera $Z = X+Y$.

1. Calcular $\text{E}(Z)$ y $\text{Var}(Z)$.
2. Calcular $P(Z = 0)$.
3. Considerar $\lambda = 1/3$ y $n = 3$. Sean Z_1, Z_2, Z_3 independientes e idénticamente distribuidas con la misma distribución que Z definida antes ($Z_1, Z_2, Z_3 \sim Z$ iid) y $W = Z_1 + Z_2 + Z_3$. Demostrar que $P(W \geq 8) < 1/4$. Sugerencia: utilizar la desigualdad de Tchebychev.
4. Se sortea una variable W y se propone el siguiente juego: si $W \geq 8$ se tiene una ganancia de 3 y en caso contrario se pierde 1. ¿Conviene jugar? Justificar.

Problema 3 (40 puntos)

1. Se considera la siguiente función:

$$f(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

   (a) Verificar que f es una función de densidad.

   (b) Calcular la función de distribución F_X de una variable aleatoria X con densidad f.

2. Se tiene una muestra iid $X_1, \ldots, X_n \sim F_X$.

   (a) Encontrar el estimador de λ por el método de los momentos.

   (b) Encontrar el estimador de λ por el método de máxima verosimilitud.

3. Se tiene la siguiente muestra de datos de tiempos de vida (en días) de mosquitos:

5,47	0,216	4,952	1,224	1,606	2,54	1,495	1,359	1,107	1,357	6,373	0,253

   (a) Estudiar la aleatoriedad de la muestra. (Realizar dos pruebas de hipótesis.)

   (b) Indicar si puede asumirse que los datos tienen distribución F_X (con F_X hallada en la parte 1.(b)), para $\lambda = 1$. (Realizar una prueba de hipótesis.)

PROBABILIDAD Y ESTADÍSTICA
SOLUCIÓN del EXAMEN, 16 de diciembre de 2010

Problema 1

1. Sea X la variable aleatoria que indica cuántas medias se sacaron hasta obtener la segunda media negra e Y la variable que indica el número de cajón elegido.

$$\begin{aligned} P(X=5) &= P(X=5|Y=1)P(Y=1)+P(X=5|Y=2)P(Y=2)+P(X=5|Y=3)P(Y=3) \\ &= \frac{1}{3}\left(0+0+P(X=5|Y=3)\right) \\ &= \frac{1}{3}\frac{2}{5}=\frac{2}{15}. \end{aligned}$$

(Sólo pueden hacerse 5 extracciones sin repetición del tercer cajón, y en ese caso la probabilidad (condicional) de tener que hacer 5 extracciones hasta obtener la segunda media negra es $\frac{3}{5}\frac{2}{4}\frac{1}{3}4=\frac{2}{5}$.)

2. El recorrido de X es $R_X=\{2,3,4,5\}$ y la función de probabilidad es la siguiente (se calcula cada caso como en la parte anterior):

$$\begin{aligned} P(X=2) &= \frac{1}{3}\left(\frac{2}{3}\frac{1}{2}+\frac{2}{4}\frac{1}{3}+\frac{2}{5}\frac{1}{4}\right)=\frac{1}{5} \\ P(X=3) &= \frac{1}{3}\left(\frac{1}{3}2+\frac{2}{4}\frac{2}{3}\frac{1}{2}2+\frac{3}{5}\frac{2}{4}\frac{1}{3}2\right)=\frac{2}{5} \\ P(X=4) &= \frac{1}{3}\left(0+\frac{2}{4}\frac{1}{3}3+\frac{3}{5}\frac{2}{4}\frac{2}{3}\frac{1}{2}\right)=\frac{4}{15} \\ P(X=5) &= \frac{2}{15} \end{aligned}$$

La función de distribución es

$$F_X(t)=\begin{cases} 0 & t<2 \\ \frac{1}{5} & 2\le t<3 \\ \frac{3}{5} & 3\le t<4 \\ \frac{13}{15} & 4\le t<5 \\ 1 & 5\le t \end{cases}$$

3. Ahora luego de elegir un cajón se extrae con reposición. Sea como antes X la variable aleatoria que indica cuántas medias se sacaron hasta obtener la segunda media negra e Y la variable que indica el número de cajón elegido. entonces $X|Y=1\sim BN(2,2/3)$, $X|Y=2\sim BN(2,1/2)$, $X|Y=3\sim BN(2,2/5)$ (siendo $BN(k,p)$ la distribución binomial negativa con parámetros k, p). Luego:

$$P(X=5) = \frac{1}{3}4\left(\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^3+\left(\frac{1}{2}\right)^5+\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^3\right)$$

Problema 2

1. Usando la linealidad del valor esperado y los valores de la esperanza de las distribuciones binomial y Poisson
$$E(Z) = E(X) + E(Y) = n\frac{1}{n} + \lambda = 1 + \lambda.$$
Como son independientes
$$Var(Z) = Var(X) + Var(Y) = n\frac{1}{n}\left(1 - \frac{1}{n}\right) + \lambda = 1 - \frac{1}{n} + \lambda.$$

2. $$P(Z = 0) = P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = \left(1 - \frac{1}{n}\right)^n e^{-\lambda}.$$

3. Para $n = 3$, $\lambda = 1/3$ se tiene $E(Z) = 4/3$ y $Var(Z) = 1$, de donde $E(W) = 4$ y $Var(W) = 3$.
$$P(W \geq 8) = P(W - 4 \geq 4) \leq P(|W - 4| \geq 4)$$
y aplicando la desigualdad de Tchebychev
$$P(W \geq 8) \leq P(|W - 4| \geq 4) \leq \frac{Var(W)}{16} = \frac{3}{16}.$$

4. Sea G la variable aleatoria que representa la ganancia del juego. Entonces $P(G = 3) = P(W \geq 8)$ y $P(G = -1) = 1 - P(W \geq 8)$.
$$E(W) = 3P(W \geq 8) - 1(1 - P(W \geq 8)) = 4P(W \geq 8) - 1 < 0,$$
de donde no conviene jugar.

Problema 3

1. (a)
$$\int_0^{+\infty} \lambda^2 x e^{-\lambda x} dx = -\lambda x e^{-\lambda x}|_0^{+\infty} + \lambda \int_0^{+\infty} e^{-\lambda x} dx = -e^{-\lambda x}|_0^{+\infty} = 1.$$
   (b) Si $t < 0$ $F_X(t) = P(X \leq t) = 0$. Si $t \geq 0$
$$\begin{aligned} F_X(t) &= \int_0^t \lambda^2 x e^{-\lambda x} dx \\ &= -\lambda x e^{-\lambda x}|_0^t + -e^{-\lambda x}|_0^t \\ &= 1 - e^{-\lambda t}(1 + \lambda t). \end{aligned}$$

2. (a) El valor esperado de X es
$$E(X) = \int_0^{+\infty} (1 - F_X(x))dx = \int_0^{+\infty} e^{-\lambda x}(1 + \lambda x)dx = \frac{2}{\lambda},$$
de donde el estimador por el método de los momentos es
$$\frac{2}{\tilde{\lambda}} = \bar{X}_n,\ \tilde{\lambda} = \frac{2}{\overline{X}_n}.$$

(b) La función de verosimilitud es

$$L(\lambda) = \prod_{i=1}^{n} \lambda^2 X_i e^{-\lambda X_i} = \lambda^{2n} \prod_{i=1}^{n} X_i e^{-\lambda X_i},$$

tomando logaritmo

$$\log(L(\lambda)) = 2n \log \lambda + \sum_{i=1}^{n} \log X_i - \lambda \sum_{i=1}^{n} X_i.$$

Derivando

$$\log(L(\lambda))\prime = \frac{2n}{\lambda} - \sum_{i=1}^{n} X_i = 0,$$

de donde el estimador es

$$\hat{\lambda} = \frac{2}{\bar{X}_n}.$$

($L(\hat{\lambda})$ es máximo por el signo de $\log(L(\lambda))\prime$.)

3. (a) Se realizan el test de rachas y el test de correlación de rangos de Spearman. Test de rachas: estadístico $R = 7$, p-valor $\alpha^* = 0,4453$, no se rechaza la hipótesis de que los datos sean iid. Test de correlación de rangos de Spearman: estadístico $r_S = 0,168$, p-valor $\alpha^* > 0,1$, no se rechaza la hipótesis de que los datos sean iid.

   (b) Se realiza el test de Kolmogorov-Smirnov. Estadístico $D_n = 0,2079$, p-valor $\alpha^* > 0,2$, no se rechaza la hipótesis de que los datos tengan la distribución F_X para $\lambda = 1$.